World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Review Article)

Check for updates

# Design of a Secure AI-Driven Adaptive Audit Transparency Engine to Improve Tax Compliance, Reduce Administrative Inefficiencies and Strengthen Overall Economic Prosperity

Faith Isabella Nayebale [1,*], Ewela Lucky Inakpenu [2], Isaac Ssambwa Makumbi [3] and Esther Makandah [3]

[1] *Maharishi International University, Fairfield, Iowa, United States of America.*
[2] *National Institute of Technology, Australia.*
[3] *University of West Georgia, Carrollton, Georgia, United States of America.*

World Journal of Advanced Research and Reviews, 2026, 29(02), 425-441

Publication history: Received on 21 December 2025; revised on 28 January 2026; accepted on 30 January 2026

Article DOI: https://doi.org/10.30574/wjarr.2026.29.2.0243

## Abstract

The increasing complexity of tax systems and the limitations of traditional rule-based audits have highlighted the need for adaptive, transparent, and efficient auditing solutions. This paper presents the design and evaluation of a Secure AI-Driven Adaptive Audit Transparency Engine (AI-AATE), a novel framework integrating machine learning, explainable AI (XAI), and human-in-the-loop oversight to enhance tax compliance, reduce administrative inefficiencies, and strengthen economic outcomes. The architecture combines supervised and unsupervised models for risk detection, continuous feedback incorporation for adaptive learning, and comprehensive audit logging to ensure transparency, fairness, and traceability. A rigorous evaluation framework employing operational Key Performance Indicators (KPIs), counterfactual simulations, and economic modeling quantifies performance across audit yield, coverage, processing efficiency, revenue recovery, and equity. Governance and trust metrics assess explainability, human oversight, and bias mitigation, linking design principles to measurable institutional outcomes. Simulation results demonstrate that AI-AATE can significantly improve detection of non-compliance, optimize resource allocation, and support equitable and accountable audit selection compared to traditional approaches. By bridging technical design, performance evaluation, and economic impact assessment, this study contributes a holistic methodology for AI-enabled audit systems, offering actionable insights for policymakers, tax authorities, and researchers. The findings underscore the potential of AI-AATE to transform public-sector auditing while maintaining fairness, legitimacy, and public trust, addressing a critical gap in the literature on adaptive, transparent, and secure AI applications in taxation.

**Keywords:** Adaptive Audit; AI-Driven Tax Compliance; Explainable AI (XAI); Governance and Transparency; Economic Impact; Public Sector Innovation

## 1. Introduction

### 1.1. Digital Transformation of Tax Administration and the Compliance Challenge

Tax administrations worldwide are undergoing a profound digital transformation driven by the expansion of electronic filing, real-time payment systems, third-party data reporting, and cross-border financial transparency initiatives (OECD, 2023). These developments have significantly increased the volume, velocity, and complexity of tax-related data available to revenue authorities. While digitalization has improved filing efficiency and reduced certain forms of evasion, it has not fully resolved persistent weaknesses in audit selection, compliance enforcement, and administrative effectiveness.

* Corresponding author: Faith Isabella Nayebale

Traditional audit systems remain constrained by limited coverage, manual workflows, and rule-based risk-scoring mechanisms that struggle to detect sophisticated or adaptive non-compliance behaviors (IMF, 2022). As a result, tax administrations frequently audit a small fraction of taxpayers, often relying on static indicators that lag behind evolving economic practices such as digital commerce, platform-based work, and complex financial engineering. These shortcomings contribute directly to the global tax gap, which continues to exceed USD 600 billion annually, largely due to underreporting, aggressive tax planning, and administrative inefficiencies (Cobham & Janský, 2023).

The persistence of this gap underscores a structural problem: while data availability has expanded dramatically, the analytical and governance frameworks required to convert data into effective, fair, and trusted enforcement decisions have not kept pace.

### 1.2. Emergence of AI-Driven Auditing and Adaptive Enforcement Models

Recent advances in artificial intelligence (AI), machine learning (ML), and large-scale analytics have created new opportunities for tax administrations to move beyond static audit selection toward adaptive, data-driven enforcement models. AI-enabled systems can process heterogeneous data sources, detect non-linear relationships, and identify anomalies that are invisible to traditional statistical methods (Zhang et al., 2022). In particular, ensemble learning, anomaly detection, and reinforcement learning techniques enable continuous refinement of risk assessments as new information becomes available.

Several tax authorities have begun experimenting with predictive analytics and AI-supported audit tools, reporting improvements in detection accuracy and administrative efficiency (OECD, 2023). However, most existing implementations focus narrowly on risk prediction or revenue maximization, with limited attention to transparency, explainability, or institutional trust. This narrow focus presents a fundamental challenge: in tax administration, audit decisions are not purely technical outputs but legally and socially consequential acts that affect taxpayer rights, perceptions of fairness, and voluntary compliance behavior (Bird & Zolt, 2022).

Without mechanisms for explainability and accountability, AI-driven audits risk reinforcing perceptions of arbitrariness or bias, potentially eroding trust and undermining the very compliance they aim to enhance.

### 1.3. Transparency, Trust, and the Governance Gap in AI-Based Auditing

Transparency has long been recognized as a cornerstone of effective tax administration and voluntary compliance. Empirical evidence suggests that taxpayers are more likely to comply when enforcement actions are perceived as fair, consistent, and procedurally just (IMF, 2023). However, many AI systems, particularly those based on complex models such as deep learning, operate as "black boxes," producing outputs that are difficult for auditors, policymakers, or taxpayers to interpret (Azmi et al., 2023).

The absence of transparent explanations for audit selection decisions creates a governance gap. While AI may improve detection rates, opaque systems can amplify legal, ethical, and reputational risks for tax authorities. Concerns over algorithmic bias, discriminatory outcomes, data misuse, and cybersecurity vulnerabilities have slowed adoption and raised questions about the legitimacy of automated enforcement (Kassa & Taibi, 2023).

Moreover, most AI applications in tax enforcement lack integrated feedback mechanisms that allow systems to learn from audit outcomes in a controlled and accountable manner. This limitation restricts their ability to adapt to emerging compliance risks while maintaining institutional oversight.

### 1.4. Toward a Secure AI-Driven Adaptive Audit Transparency Engine

Addressing these challenges requires a shift from isolated AI tools toward holistic, governance-aware architectures that integrate adaptive analytics with transparency, security, and performance monitoring. This study proposes the design of a Secure AI-Driven Adaptive Audit Transparency Engine (AATE) as a response to this unmet need.

The AATE concept extends beyond conventional risk-scoring systems by embedding explainable AI (XAI), real-time audit logging, cybersecurity safeguards, and performance dashboards into a unified framework. Rather than treating transparency as an afterthought, the engine is designed to generate human-interpretable explanations for audit decisions, enabling auditors to understand, validate, and communicate AI-driven outcomes. At the same time, adaptive learning mechanisms allow the system to refine audit priorities as taxpayer behaviors evolve, while maintaining institutional control through human-in-the-loop oversight. By aligning technological capability with governance principles, the AATE framework seeks to reconcile efficiency gains with fairness, accountability, and trust.

## 2. Background and Literature Review

### 2.1. Digital Transformation of Tax Administration and the Rise of Intelligent Auditing

Over the past two decades, tax administrations worldwide have undergone significant digital transformation, driven by the expansion of electronic filing systems, third-party information reporting, and digital payment infrastructures (OECD, 2023). These developments have fundamentally altered the scale, velocity, and complexity of tax data, creating both opportunities and challenges for compliance enforcement. While digitalization has increased reporting coverage and reduced manual errors, it has also exposed the limitations of traditional audit frameworks that rely heavily on static rules, manual reviews, and ex post enforcement (IMF, 2022).

Artificial intelligence (AI) and machine learning (ML) have emerged as critical tools for addressing these limitations. By enabling automated pattern recognition across high-dimensional datasets, AI systems allow tax authorities to identify hidden relationships, detect anomalies, and prioritize enforcement actions more effectively than conventional methods (Edupuganti, 2024). Unlike deterministic rule-based systems, ML models can learn from historical outcomes and continuously refine their predictions, making them particularly well suited for environments characterized by evolving taxpayer behavior and strategic non-compliance (Salmanov, 2024).

Empirical studies indicate that AI-enabled audit systems significantly outperform traditional approaches in detecting underreporting, fraudulent claims, and aggressive tax planning (Shehu & Olukeye, 2024). International bodies such as the OECD report increasing adoption of predictive analytics, network analysis, and anomaly detection models within advanced tax administrations, particularly for VAT fraud, transfer pricing risk assessment, and large-taxpayer compliance monitoring (OECD, 2025). The experience of Austria's Predictive Analytics Competence Centre, which processes millions of tax records annually using ML-based risk scoring, illustrates the scalability and practical viability of AI-driven audit selection (OECD, 2025).

Despite these advances, the literature emphasizes that technological capability alone is insufficient. The effectiveness of AI in tax administration depends on institutional context, governance frameworks, and the extent to which automated decisions are transparent and accountable (Bird & Zolt, 2022).

### 2.2. Continuous Auditing and the Shift from Episodic to Adaptive Enforcement

Traditional tax audits are inherently episodic, retrospective, and resource constrained. Typically, only a small fraction of taxpayers are audited in any given cycle, leaving substantial non-compliance undetected and creating weak deterrence effects (Alles et al., 2022). Continuous auditing (CA) represents a paradigm shift, moving from periodic inspection to ongoing monitoring supported by real-time or near real-time data flows.

The integration of AI with continuous auditing enables dynamic risk assessment, allowing audit priorities to be adjusted as new information becomes available (Alles et al., 2022). Rather than relying on fixed thresholds or static risk indicators, adaptive systems can respond to emerging patterns such as sudden revenue drops, abnormal transaction networks, or changes in filing behavior (Edupuganti, 2024).

However, existing literature highlights that most implementations of continuous auditing remain fragmented and experimental. In particular, the concept of continuous auditing of AI systems themselves (CAAI), monitoring model performance, bias, and drift over time, has received limited attention (Iskandarova et al., 2022). This gap is especially problematic in tax administration, where opaque or poorly governed AI systems can undermine procedural fairness and erode public trust.

Recent econometric evidence demonstrates that ML-based audit selection substantially increases expected revenue recovery compared to random or heuristic selection methods (Refining Public Policies with Machine Learning, 2024). Yet, these studies largely focus on revenue outcomes, offering limited insight into transparency, governance, or long-term compliance behavior.

### 2.3. Transparency, Explainability, and Trust in AI-Based Auditing

#### *2.3.1. Data Quality, Integration, and Structural Complexity*

AI-driven audit systems depend on the integration of heterogeneous data sources, including tax returns, financial statements, banking transactions, customs records, and third-party reports. While data integration enhances analytical power, it also introduces risks related to data inconsistency, duplication, and measurement error (Salmanov, 2024).

Weak data governance can propagate errors through ML models, resulting in distorted risk scores and inefficient audit targeting (Shehu & Olukeye, 2024).

Moreover, the increasing use of unstructured data raises methodological challenges related to natural language processing, data labeling, and validation (Transforming Auditing in the AI Era, 2025).

#### *2.3.2. Explainability and Algorithmic Accountability*

A central concern in AI-based tax auditing is explainability. Complex ML models, particularly ensemble and deep learning architectures, often function as "black boxes," producing accurate predictions without transparent reasoning (Alles et al., 2022). In high-stakes regulatory environments, such opacity can conflict with principles of due process, accountability, and the right to explanation.

Explainable AI (XAI) techniques, such as SHAP and LIME, have been proposed as partial solutions, enabling feature-level interpretation of model outputs (Ribeiro et al., 2016). However, the literature notes that explanation alone does not guarantee fairness or trust; explanations must be intelligible, auditable, and embedded within institutional governance structures (Transforming Auditing in the AI Era, 2025).

#### *2.3.3. Organizational and Regulatory Constraints*

Beyond technical challenges, AI adoption in tax auditing is constrained by organizational capacity, skills shortages, and regulatory uncertainty. Successful deployment requires not only data scientists and engineers but also auditors capable of interpreting AI outputs and exercising informed judgment (Iskandarova et al., 2022). Regulatory frameworks in many jurisdictions have yet to fully address liability, accountability, and appeal mechanisms for AI-generated audit decisions.

### 2.4. Empirical Evidence on AI-Supported Audit Performance

Recent empirical studies provide growing evidence of the tangible benefits of AI-supported auditing. Hybrid models combining supervised and unsupervised learning consistently outperform single-method approaches in detecting anomalies while reducing manual workload (Transforming Auditing in the AI Era, 2025). Organizations adopting AI, robotic process automation (RPA), and natural language processing (NLP) report reductions in compliance time of 30–60% and error rates of 35–45% (Shehu & Olukeye, 2024).

Systematic reviews further indicate that AI-enabled auditing is now the most prominent theme in contemporary accounting and auditing research, reflecting both academic and policy interest (Iskandarova et al., 2022). Nonetheless, data security, bias, and governance remain persistent concerns, particularly in public-sector applications. These findings suggest that while AI technologies are mature enough for operational use, their institutional integration remains incomplete.

### 2.5. Synthesis and Research Gaps

The reviewed literature reveals several unresolved gaps. First, few studies propose end-to-end architectures that integrate adaptive risk scoring, continuous monitoring, explainable decision-making, and secure audit logging within a single system (Alles et al., 2022). Second, empirical research rarely examines long-term effects on voluntary compliance, administrative efficiency, and public trust (Refining Public Policies with Machine Learning, 2024). Third, governance mechanisms for auditing the auditors remain underdeveloped (Transforming Auditing in the AI Era, 2025).

These gaps motivate the need for a Secure AI-Driven Adaptive Audit Transparency Engine (AI-AATE) that explicitly embeds transparency, explainability, cybersecurity, and governance into its core design rather than treating them as secondary considerations.

**Table 1** Summary of Key Studies on AI and Machine Learning in Auditing and Tax Compliance

| Year | Author(s) | Context / Domain | Methodology | Main Findings | Limitations |
|---|---|---|---|---|---|
| 2024 | Edupuganti | Fraud detection & compliance | Review of AI/ML applications in auditing | AI/ML improves anomaly detection and automates data processing, freeing | Limited empirical validation in large-scale tax administrations |

| | | | | | |
|---|---|---|---|---|---|
| | | | | auditors for judgmental tasks | |
| 2024 | Shehu & Olukeye | Tax compliance reporting | Case studies & data analysis | AI reduces reporting errors, shortens compliance time, and supports regulatory functions | Focused mainly on corporate filings; lacks governance/ethical analysis |
| 2022 | Alles, Brennan, Kogan, & Vasarhelyi | Continuous auditing | Literature review + conceptual framework | Continuous auditing combined with AI can enhance audit coverage and timeliness | Few large-scale implementations; governance challenges remain |
| 2024 | Salmanov | Corporate governance & fraud | Empirical study with ML models | ML tools improve fraud detection accuracy and optimize audit focus | Dataset limited to select firms; not generalized to tax authorities |
| 2025 | OECD | Government tax administration | Policy review & case studies | AI/ML adoption improves audit selection and efficiency; Austria example shows revenue gains | Mainly descriptive; lacks detailed design frameworks |
| 2024 | "Refining public policies with ML" | Tax auditing | Empirical econometric study | ML-based audit selection increases expected revenue recovery | Focused on revenue outcomes; limited coverage of transparency and trust issues |
| 2025 | Transforming Auditing in AI Era | Accounting & auditing | Systematic review of 465 papers | AI adoption reduces compliance time and errors; highlights key thematic areas | Limited discussion of governance-aware architectures; mostly secondary studies |
| 2022 | Iskandarova, Jones, & Li | Audit optimization | Systematic review | AI/ML adoption improves efficiency and anomaly detection; supports audit decisions | Few practical implementations; lacks integration with continuous auditing models |

To summarize prior empirical and theoretical work, Table 2.1 presents key studies on AI and machine learning in auditing and tax compliance, highlighting their context, methodology, main findings, and limitations. The table underscores the consistent benefits of AI for anomaly detection, audit efficiency, error reduction, and revenue improvement. However, it also reveals persistent gaps: most studies focus on narrow applications, lack integrated architectures for continuous auditing, and provide limited guidance on governance, transparency, and trust mechanisms. This synthesis provides a strong rationale for designing a holistic, secure, and adaptive audit transparency engine, which the subsequent chapter proposes.

## 3. Design and Architecture of the AI-AATE

### 3.1. Design Objectives and System Requirements

The design of the Secure AI-Driven Adaptive Audit Transparency Engine (AI-AATE) is grounded in the theoretical gaps and practical limitations identified in prior research on AI-enabled auditing and tax administration (Alles et al., 2022; OECD, 2025). Existing systems often emphasize predictive accuracy or revenue recovery while underemphasizing transparency, governance, cybersecurity, and long-term institutional trust. As a result, many AI-based audit tools remain fragmented, opaque, or difficult to justify within legal and administrative frameworks.

To address these shortcomings, AI-AATE is conceived as a governance-aware, adaptive audit system that balances analytical sophistication with explainability, accountability, and security. Rather than optimizing a single objective (e.g., detection accuracy), the system is designed to satisfy a multidimensional set of requirements reflecting the realities of public-sector enforcement.

#### *3.1.1. Core Design Objectives*

The primary objective of AI-AATE is to enhance audit effectiveness while preserving procedural fairness and institutional legitimacy. This overarching goal is decomposed into the following design objectives:

- Adaptive Risk Detection: Enable continuous learning from evolving taxpayer behavior through machine learning models capable of updating risk assessments dynamically as new data and audit outcomes become available.
- Transparency and Explainability: Ensure that audit selection decisions are interpretable by auditors, policymakers, and taxpayers, supporting explainability, appealability, and accountability.
- Operational Efficiency: Improve audit coverage and resource allocation by prioritizing high-risk cases while reducing unnecessary audits of compliant taxpayers.
- Security and Privacy Preservation: Protect sensitive taxpayer data and audit intelligence through robust cybersecurity controls, privacy-by-design principles, and secure access mechanisms.
- Governance and Human Oversight: Embed human-in-the-loop controls, decision checkpoints, and audit-of-audit mechanisms to prevent unchecked automation and algorithmic drift.

These objectives reflect a shift from narrowly defined AI performance metrics toward a broader conception of audit system quality, consistent with emerging guidance on responsible AI in public administration (OECD, 2025).

#### *3.1.2. Functional Requirements*

To operationalize the design objectives, AI-AATE must satisfy a set of functional requirements that define what the system must do:

- Multi-Source Data Ingestion: Seamlessly integrate structured and unstructured data from tax filings, financial transactions, third-party reports, and historical audit records.
- Risk Scoring and Prioritization: Generate probabilistic risk scores for each taxpayer or filing, enabling ranked audit selection under resource constraints.
- Anomaly Detection: Identify deviations from expected behavior patterns using supervised and unsupervised learning techniques.
- Continuous Learning: Update models based on audit outcomes, feedback from auditors, and changes in economic conditions.
- Explainable Outputs: Produce feature-level explanations and audit rationales that accompany each audit recommendation.
- Performance Monitoring: Track system-level KPIs such as detection accuracy, audit yield, false-positive rates, and processing latency.

#### *3.1.3. Non-Functional Requirements*

- Beyond functionality, AI-AATE must meet several non-functional requirements critical to its sustainability and legitimacy:
- Scalability: Support large taxpayer populations and high-frequency data streams without degradation in performance.
- Reliability and Robustness: Maintain stable operation under data noise, partial outages, or adversarial behavior.
- Interoperability: Integrate with existing tax administration systems, case management platforms, and reporting tools.
- Maintainability: Allow for modular updates to models, rules, and governance policies without system-wide redesign.

These requirements address concerns raised in the literature regarding the fragility and inflexibility of many AI deployments in public-sector contexts (Iskandarova et al., 2022).

*3.1.4. Governance and Ethical Requirements*

A defining feature of AI-AATE is the explicit treatment of governance and ethics as design-time constraints, not post-deployment considerations. Accordingly, the system incorporates the following governance requirements:

- Human-in-the-Loop Controls: Final audit decisions remain subject to human review, particularly for high-risk or high-impact cases.
- Auditability of AI Decisions: All model outputs, explanations, and overrides are logged in tamper-resistant records to support internal and external review.
- Bias Monitoring and Mitigation: Continuous assessment of disparate impacts across taxpayer segments, with corrective mechanisms where necessary.
- Appeal and Review Mechanisms: Support structured review processes for contested audit decisions, grounded in explainable system outputs.

These requirements directly respond to concerns about algorithmic bias, opacity, and accountability highlighted in prior studies on AI-based auditing (Azmi et al., 2023; Alles et al., 2022).

### 3.2. Conceptual Architecture of the AI-Driven Adaptive Audit Transparency Engine (AI-AATE)

The AI-Driven Adaptive Audit Transparency Engine (AI-AATE) is designed as a layered, modular architecture that operationalizes the design objectives and system requirements established in Section 3.1. Rather than functioning as a monolithic risk-scoring tool, the architecture emphasizes separation of concerns, enabling adaptability, transparency, and governance to be embedded directly into the system's structure. This architectural approach reflects best practices in responsible AI system design for public-sector applications (OECD, 2025; Alles et al., 2022).

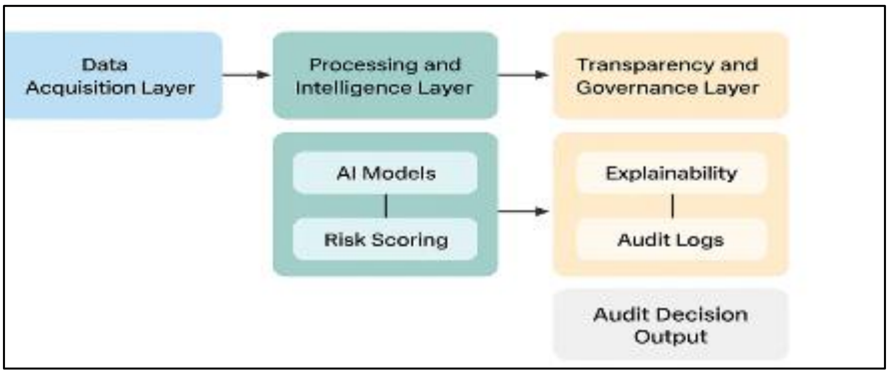


**Figure 1** AI-AATE System Architecture and Data Flow

Figure 1 illustrates the high-level architecture of AI-AATE, showing the flow of information from data acquisition through AI processing and governance layers to final audit decision outputs. Each layer performs a distinct role while remaining tightly integrated through secure interfaces and feedback mechanisms.

*3.2.1. Data Acquisition and Integration Layer*

At the foundation of AI-AATE is the data acquisition and integration layer, which consolidates heterogeneous data sources into a unified analytical environment. Modern tax administrations generate vast volumes of data from electronic filings, e-invoicing systems, third-party reports, financial institutions, customs records, and historical audit outcomes. However, these data are often fragmented, inconsistently formatted, and subject to quality issues (Salmanov, 2024).

This layer performs data ingestion, validation, standardization, and enrichment, ensuring that downstream AI models operate on consistent and reliable inputs. Structured data (e.g., declared income, VAT filings) are integrated alongside semi-structured and unstructured sources (e.g., transaction narratives, supporting documents), reflecting the growing complexity of taxpayer behavior in digital economies (Shehu & Olukeye, 2024).

Crucially, the architecture treats data governance as a core architectural concern, embedding access controls, encryption, and provenance tracking at the ingestion stage. This design choice reduces downstream privacy risks and supports regulatory compliance from the outset (OECD, 2023).

#### *3.2.2. AI Analytics and Adaptive Risk Modeling Layer*

The AI analytics layer constitutes the analytical core of AI-AATE. It combines supervised learning, unsupervised anomaly detection, and ensemble techniques to generate probabilistic risk assessments for audit selection. Unlike static rule-based systems, this layer continuously adapts by incorporating new data, audit results, and feedback from human auditors.

Supervised models leverage labeled audit outcomes to predict the likelihood of non-compliance, while unsupervised techniques identify emerging patterns and deviations not captured by historical labels (Edupuganti, 2024). Ensemble approaches balance predictive accuracy with robustness, mitigating the risk of overfitting or reliance on a single model type (Transforming Auditing in the AI Era, 2025).

Adaptivity is achieved through periodic retraining and performance monitoring, enabling the system to respond to evolving taxpayer strategies and macroeconomic changes. This capability directly addresses limitations identified in the literature regarding the rigidity of traditional audit selection frameworks (OECD, 2021; Alles et al., 2022).

#### *3.2.3. Explainability and Transparency Layer*

A defining feature of AI-AATE is the explicit separation of explainability and transparency functions from core analytics. While complex AI models may be required for accurate risk detection, their outputs are systematically translated into interpretable explanations through an intermediate transparency layer.

This layer generates feature-level contributions, rule-based approximations, and narrative explanations that clarify why a particular taxpayer or transaction has been flagged for audit. Such explanations are essential for internal accountability, legal defensibility, and taxpayer trust in automated enforcement systems (Azmi et al., 2023).

By architecturally isolating explainability mechanisms, AI-AATE avoids the common trade-off between model performance and interpretability, instead treating transparency as a non-negotiable system output rather than an optional add-on (Iskandarova et al., 2022).

#### *3.2.4. Governance, Oversight, and Audit-of-Audit Layer*

Above the analytics and transparency layers sits the governance and oversight layer, which ensures that AI-driven recommendations remain subject to institutional control. This layer enforces human-in-the-loop decision points, particularly for high-risk or high-impact cases, preventing unchecked automation.

All model outputs, explanations, overrides, and final decisions are logged in tamper-resistant audit trails, enabling retrospective review and external accountability. This “audit-of-audit” capability responds directly to concerns about algorithmic opacity and regulatory legitimacy raised in the literature (Alles et al., 2022; OECD, 2025).

In addition, governance modules continuously monitor bias indicators, model drift, and compliance with predefined ethical and legal constraints. This ensures that AI-AATE evolves within acceptable institutional boundaries rather than optimizing narrowly defined technical metrics.

#### *3.2.5. Audit Decision and Feedback Layer*

The final layer translates AI-AATE outputs into operational audit decisions. Risk-ranked cases are forwarded to audit teams along with accompanying explanations, confidence scores, and relevant supporting evidence. Importantly, the system captures feedback from audit outcomes which is fed back into the analytics layer to support continuous learning.

This closed-loop design reinforces both adaptivity and accountability, aligning system behavior with real-world enforcement outcomes rather than static assumptions (Refining Public Policies with Machine Learning, 2024).

### 3.3. AI Models and Algorithms for Adaptive Audit Selection

The effectiveness of the AI-Driven Adaptive Audit Transparency Engine (AI-AATE) depends critically on the design of its analytical core. This section describes the AI models, feature representations, and adaptive learning mechanisms used to generate risk-based audit recommendations. Rather than relying on a single predictive technique, AI-AATE adopts a hybrid and ensemble-based modeling strategy to balance accuracy, robustness, and explainability

#### *3.3.1. Feature Engineering and Risk Signal Construction*

Audit risk prediction requires the transformation of raw administrative data into meaningful risk signals that capture financial inconsistencies, behavioral anomalies, and structural relationships. AI-AATE employs multi-dimensional feature engineering across four primary categories:

Financial Features capture discrepancies between reported income, expenses, and third-party information. Examples include abnormal profit margins, VAT input-output mismatches, unexplained revenue volatility, and deviations from sectoral benchmarks (Salmanov, 2024).

Behavioral Features model filing behavior over time, including late submissions, frequent amendments, irregular payment patterns, and sudden changes in declared activity. Temporal sequencing of such behaviors has been shown to improve early detection of non-compliance (Edupuganti, 2024).

Relational and Network Features exploit links between taxpayers, suppliers, and counterparties. Network centrality, transaction clustering, and shared identifiers can reveal coordinated underreporting or carousel fraud patterns that are invisible to entity-level analysis (OECD, 2025).

Macroeconomic and Contextual Features adjust risk assessments for sector-specific trends, regional economic shocks, and regulatory changes, reducing false positives during periods of legitimate economic disruption (IMF, 2023).

#### *3.3.2. Supervised Risk Prediction Models*

Supervised learning models form the primary predictive backbone of AI-AATE, using historical audit outcomes as labeled data. Gradient boosting machines (GBM), random forests (RF), and regularized logistic regression models are particularly well-suited for this task due to their ability to handle non-linear relationships and heterogeneous features (Shehu & Olukeye, 2024).

Let

$$X_i = (x_{i1}, x_{i2}, \ldots, x_{ik}) \qquad 3.1$$

represent the feature vector for taxpayer *i*, and

$$y_i \in \{0,1\} \qquad 3.2$$

denote observed audit outcomes (non-compliance detected or not). The supervised model estimates:

$$P(y_i = 1 \mid X_i) \qquad 3.3$$

which is interpreted as the probabilistic audit risk score.

These models are evaluated using precision-recall metrics rather than accuracy alone, reflecting the highly imbalanced nature of audit populations (Alles et al., 2022).

#### *3.3.3. Unsupervised Anomaly Detection for Emerging Risks*

Because historical audit data may not capture newly emerging compliance strategies, AI-AATE incorporates unsupervised anomaly detection to identify novel patterns. Techniques such as Isolation Forests and autoencoder-based reconstruction error models flag observations that deviate significantly from learned norms (Transforming Auditing in the AI Era, 2025).

Unsupervised scores do not directly trigger audits but serve as early warning signals, prompting closer review or model recalibration. This dual-track design prevents over-reliance on historical labels and improves resilience to strategic adaptation by taxpayers.

#### *3.3.4. Ensemble Risk Scoring and Decision Logic*

To integrate insights from multiple models, AI-AATE applies an ensemble aggregation function:

$$R_i = \sum_{m=1}^{M} w_m \cdot r_{im} \quad 3.4$$

where $r_{im}$ is the risk score from model *m* and $w_m$ represents dynamically adjusted model weights.

Weights are calibrated based on recent predictive performance, stability, and explainability metrics. This ensemble approach improves robustness and reduces the likelihood that any single modeling assumption disproportionately influences audit selection (Edupuganti, 2024).

Final audit prioritization incorporates resource constraints and risk thresholds, ensuring that audit capacity is allocated to cases with the highest expected revenue impact.

*3.3.5. Adaptive Learning and Feedback Mechanisms*

Adaptivity is a defining characteristic of AI-AATE. Audit outcomes, auditor feedback, and appeal results are continuously fed back into the learning pipeline. Model drift is monitored using statistical divergence measures, triggering retraining when predictive performance degrades or data distributions shift (Iskandarova et al., 2022).

This feedback-driven learning cycle aligns the system with evolving economic behavior while maintaining institutional oversight, directly addressing critiques of static and opaque audit selection systems in prior literature (OECD, 2021).

### 3.4. Explainability, Transparency, and Auditability Mechanisms

The deployment of AI in tax audit selection introduces significant legitimacy, legal, and ethical considerations. In high-stakes regulatory environments, predictive accuracy alone is insufficient; audit decisions must also be explainable, traceable, and contestable. This section describes the mechanisms through which AI-AATE embeds explainability, transparency, and auditability as core system properties, rather than auxiliary features.

*3.4.1. Explainable AI Layer for Audit Decisions*

AI-AATE integrates an explicit Explainable AI (XAI) layer that operates alongside the ensemble risk models described in Section 3.3. This layer translates complex model outputs into interpretable explanations that identify the most influential risk factors contributing to each audit recommendation.

Feature attribution methods are used to quantify the contribution of individual features to a taxpayer's risk score.

For a given taxpayer *i*, the explanation function can be expressed as:

$$R_i = \phi_0 + \sum_{j=1}^{k} \phi_{ij} \quad 3.5$$

where $\phi_{ij}$ represents the marginal contribution of feature *j* to the overall risk score. These explanations allow auditors to understand why a case was selected, not merely that it was selected (Iskandarova et al., 2022).

Importantly, explanations are generated at multiple levels of abstraction: technical explanations for analysts and simplified narratives for oversight bodies and taxpayers. This multi-tiered explainability mitigates the risk of misinterpretation while preserving analytical rigor (Azmi et al., 2023).

*3.4.2. Transparency of Audit Selection Logic*

Transparency within AI-AATE extends beyond individual explanations to encompass the audit selection process as a whole. The system maintains documented selection criteria, model versions, and risk thresholds used at each decision point. This ensures that audit practices remain consistent over time and can be externally reviewed if required.

Rather than exposing proprietary model internals, AI-AATE adopts a procedural transparency approach, whereby stakeholders can verify that decisions were made according to predefined, lawful, and nondiscriminatory rules (OECD, 2025). This approach balances the need for accountability with the practical necessity of protecting sensitive enforcement methodologies.

#### *3.4.3. Audit-of-Audit and Decision Traceability*

To support institutional accountability, AI-AATE implements an audit-of-audit framework in which all stages of the audit decision lifecycle are logged in immutable, time-stamped records. These records capture input data versions, model outputs, explanation artifacts, human overrides, and final audit outcomes.

Such traceability enables retrospective evaluation of system behavior, facilitates internal audits, and supports judicial or parliamentary oversight where necessary (Alles et al., 2022). By preserving a complete decision trail, AI-AATE reduces institutional risk and strengthens public trust in automated enforcement systems.

#### *3.4.4. Human Oversight, Override, and Contestability*

Despite its advanced automation capabilities, AI-AATE is explicitly designed to preserve human authority over audit decisions. High-risk cases, borderline scores, and socially sensitive scenarios trigger mandatory human review before enforcement actions proceed.

Auditors retain the ability to override system recommendations, with all overrides recorded and analyzed to identify systematic model limitations or training gaps. Furthermore, the transparency layer enables contestability, allowing taxpayers to challenge audit decisions using documented explanations and evidence trails (Bird & Zolt, 2022).

### 3.5. Governance, Transparency, and Operational Integrity

AI-AATE embeds explainability, transparency, and governance within its core architecture to ensure technical performance aligns with institutional credibility. An explainable AI layer translates risk scores into interpretable feature contributions and narrative summaries, enabling auditors and oversight bodies to understand why specific cases are selected. Procedural transparency is reinforced through version-controlled model logs, recorded thresholds, and human intervention records, providing an auditable trail that supports accountability and regulatory compliance.

Operational integrity is maintained via a controlled workflow: data acquisition, feature generation, risk modeling, and audit prioritization occur in modular, monitored stages. Feedback from audit outcomes is systematically fed back into adaptive models. Human-in-the-loop checkpoints for high-risk or sensitive cases preserve oversight and contestability, ensuring fairness and trust.

By integrating these mechanisms, AI-AATE demonstrates that advanced, adaptive audit intelligence can be both technically rigorous and institutionally trustworthy, while enabling evaluation through measurable KPIs for efficiency, fairness, and transparency.

### 3.6. Conceptual Workflow and Adaptive Operation

AI-AATE operates as a modular, adaptive audit system in which data acquisition, feature generation, risk modeling, audit prioritization, and feedback integration are executed in a controlled pipeline. Each stage is designed to maintain data integrity, traceability, and institutional oversight, ensuring that no decision proceeds without appropriate checks.

Risk modeling combines supervised predictions with unsupervised anomaly detection, producing a probabilistic risk score for each taxpayer. High-risk or borderline cases are routed through human-in-the-loop review to preserve accountability, fairness, and contestability. Feedback from audit outcomes is systematically incorporated into model retraining and threshold adjustments, allowing the system to adapt to evolving compliance patterns.

The workflow emphasizes auditability and transparency: all key operations, decisions, and human interventions are logged for retrospective review. This enables evaluation against KPIs for efficiency, coverage, and fairness without revealing proprietary algorithms or configuration parameters. A conceptual process diagram (Figure 1) summarizes these stages, illustrating the continuous flow from data acquisition through adaptive decision-making to feedback integration.

## 4. Performance Evaluation, KPIs, and Economic Impact Modeling

### 4.1. Evaluation Philosophy and Baseline Definition

Evaluating AI-driven audit systems presents inherent methodological challenges, particularly in the absence of full-scale operational deployment. Unlike conventional information systems, audit engines directly influence taxpayer behavior, enforcement outcomes, and institutional legitimacy. Consequently, evaluation must extend beyond predictive accuracy to encompass administrative efficiency, economic impact, fairness, and governance robustness. This chapter adopts a design science–oriented evaluation philosophy, emphasizing *ex-ante* assess ability, counterfactual reasoning, and scenario-based analysis rather than purely retrospective performance measurement (Hevner et al., 2004; OECD, 2025).

#### *4.1.1. Evaluation Objectives and Scope*

The primary objective of the evaluation framework is to assess whether AI-AATE can outperform traditional audit selection mechanisms while preserving transparency, accountability, and institutional trust. Specifically, the evaluation framework addresses four interrelated dimensions:

- Effectiveness – the system's ability to identify high-risk cases and improve audit yield.
- Efficiency – reductions in administrative workload and audit costs per unit of revenue recovered.
- Equity and Governance – fairness, explainability, and contestability of audit decisions.
- Economic Impact – effects on revenue mobilization, compliance behavior, and macroeconomic stability.

This multidimensional scope reflects consensus in the literature that AI systems in public administration must be evaluated against broader societal objectives rather than narrow technical benchmarks (Bird & Zolt, 2022; IMF, 2023).

#### *4.1.2. Baseline Audit Selection Framework*

To enable counterfactual comparison, AI-AATE is evaluated against a baseline audit selection framework representative of prevailing practices in many tax administrations. The baseline system is characterized by:

- Rule-based risk scoring using static thresholds
- Limited use of historical audit outcomes for learning
- Periodic (rather than continuous) audit cycles
- Minimal transparency regarding selection logic
- Manual intervention at late stages of audit selection

Formally, the baseline risk score for taxpayer *i* can be expressed as:

$$R_i^{base} = \sum_{j=1}^{k} \alpha_j . x_{ij} \qquad 4.1$$

where $x_{ij}$ represents predefined risk indicators and $\alpha_j$ are fixed weights determined through expert judgment or legacy policy rules.

This formulation contrasts with the adaptive, ensemble-based risk estimation function defined in Chapter 3, highlighting the structural limitations of static audit frameworks (OECD, 2021).

#### *4.1.3. Counterfactual Evaluation Logic*

Because real-world audit outcomes are only observed for selected cases, direct comparison between AI-AATE and baseline systems requires counterfactual reasoning. The evaluation framework therefore relies on parallel scoring and simulation, where both systems score the same taxpayer population under identical constraints.

Let $S^{AI}$ and $S^{base}$ denote the sets of taxpayers selected for audit by AI-AATE and the baseline system, respectively, given equal audit capacity $C$. Differences in outcomes are evaluated using expected values rather than realized outcomes:

$$\Delta E(Y) = \mathbf{Z}[Y \mid S^{AI}] - \mathbf{Z}[Y \mid S^{base}] \qquad 4.2$$

where $Y$ represents outcome variables such as detected non-compliance, recovered revenue, or audit duration. This approach aligns with established evaluation methods for policy algorithms where randomized experimentation is infeasible (Refining Public Policies with Machine Learning, 2024).

#### *4.1.4. Evaluation Time Horizon and Learning Effects*

The evaluation explicitly accounts for dynamic learning effects. Unlike static systems, AI-AATE's performance evolves over time as models adapt to new data and behavioral responses. Evaluation is therefore conducted across multiple simulated audit cycles, capturing:

- Short-term performance gains
- Medium-term learning improvements
- Long-term stabilization or saturation effects

This temporal perspective prevents overestimation of early gains and enables realistic assessment of sustainability (Iskandarova et al., 2022).

#### *4.1.5. Constraints and Assumptions*

To ensure transparency and replicability, the evaluation framework operates under clearly stated constraints:

- Audit capacity is fixed and equal across systems
- Legal and procedural rules are held constant
- No behavioral deterrence effects are assumed unless explicitly modeled
- Data quality limitations are explicitly parameterized

These assumptions allow the evaluation to isolate the incremental value of adaptivity, transparency, and governance, rather than conflating system design with external policy changes (IMF, 2023).

### **4.2. Operational KPIs and Performance Metrics**

To evaluate AI-AATE's effectiveness, efficiency, and governance, a set of quantifiable Key Performance Indicators (KPIs) is defined. These KPIs translate the system's architecture and workflow into measurable outputs that can be simulated or assessed against baseline audit frameworks.

#### *4.2.1. Effectiveness Metrics*

Audit Yield ( $Y$ ): Proportion of audits detecting confirmed non-compliance.

$$Y = \frac{\text{Number of confirmed non-compliant cases}}{\text{Total audits conducted}} \tag{4.3}$$

Coverage ( $C$ ): Fraction of taxpayer population effectively assessed by the system.

$$C = \frac{\text{Number of taxpayers evaluated}}{\text{Total taxpayer population}} \tag{4.4}$$

Detection Accuracy ( $DA$ ): Probability that high-risk taxpayers are correctly flagged.

$$DA = \frac{\text{True Positives}}{\text{True Positives + False negatives}} \tag{4.5}$$

#### *4.2.2. Efficiency Metrics*

Audit Efficiency ( $AE$ ): Revenue recovered per unit of administrative effort.

$$AE = \frac{\text{Revenue recovered}}{\text{Auditor hours or cost}} \tag{4.6}$$

Processing Time ( $PT$ ): Average time to evaluate a taxpayer case. Lower values indicate streamlined operations.

#### *4.2.3. Governance and Fairness Metrics*

Explainability Score ( $EX$ ): Proportion of audit decisions accompanied by interpretable rationale.

$$EX = \frac{\text{Audits with full explanation}}{\text{Total audits}} \tag{4.7}$$

- Human Override Rate ( $HO$ ): Fraction of automated recommendations adjusted by human auditors, indicating model conservatism and oversight engagement.
- Equity Index ( $EI$ ): Assesses disparity in audit selection across demographic or economic strata, highlighting bias potential.

#### *4.2.4. Reliability and Adaptivity Metrics*

- Model Drift Rate ( $MD$ ): Frequency and magnitude of retraining triggered by changing data patterns.
- Feedback Incorporation ( $FI$ ): Fraction of audit outcomes integrated into subsequent model updates, reflecting adaptive learning.

### **4.3. Simulation and Counterfactual Modeling**

To evaluate AI-AATE without full-scale deployment, a simulation-based counterfactual framework is employed. This approach compares the adaptive AI system against a baseline rule-based audit framework, holding audit capacity and procedural rules constant. The simulation assesses expected outcomes across effectiveness, efficiency, and governance dimensions.

#### *4.3.1. Counterfactual Logic*

Let $S^{AI}$ and $S^{base}$ denote the sets of taxpayers selected for audit by AI-AATE and the baseline system, respectively. For any outcome variable $Y$ (e.g., detected non-compliance, revenue recovered), the expected performance gain is:

$$\Delta E(Y) = \mathbf{Z}[Y \mid S^{AI}] - \mathbf{Z}[Y \mid S^{base}] \tag{4.8}$$

This formulation isolates the incremental contribution of adaptive, transparent audit selection while controlling for external variables.

#### *4.3.2. Simulation Design*

The simulation replicates multiple audit cycles to capture short-term, medium-term, and long-term performance:

- Short-term: Immediate improvements in detection and coverage.
- Medium-term: Adaptation through model retraining based on audit outcomes.
- Long-term: Stabilization of efficiency, fairness, and compliance improvements.

Key inputs include historical audit data, taxpayer risk indicators, and procedural rules. Monte Carlo or agent-based simulation methods can be used to model variability and uncertainty in compliance behavior.

#### *4.3.3. Linking to KPIs*

Simulation outputs are directly mapped to the KPIs defined in Section 4.2:

- Audit yield ( $Y$ ) and coverage ( $C$ ) measure effectiveness.
- Audit efficiency ( $AE$ ) and processing time ( $PT$ ) measure operational efficiency.
- Explainability ( $EX$ ), human override ( $HO$ ), and equity index ( $EI$ ) measure governance and fairness.
- Model drift ( $MD$ ) and feedback incorporation ( $FI$ ) measure adaptivity and reliability.

### 4.4. Economic Impact and Cost–Benefit Analysis

AI-AATE's potential value extends beyond operational performance to economic and fiscal outcomes. This section models the system's impact on revenue mobilization, administrative efficiency, and broader compliance behavior, integrating the results of KPIs and simulation outputs.

#### *4.4.1. Revenue Mobilization*

Expected revenue gains are computed as the difference in audit yield between AI-AATE and baseline systems:

$$\Delta R = \sum_{i \in S^{AI}} E[Tax_i] - \sum_{i \in S^{base}} E[Tax_i] \qquad 4.9$$

where $E[Tax_i]$ represents the estimated recoverable tax for taxpayer *i*. Simulated audit cycles capture short- and medium-term effects, including improvements due to adaptive learning and model retraining.

#### *4.4.2. Administrative Efficiency and Cost Reduction*

Operational efficiency gains are assessed by comparing recovered revenue per auditor hour between AI-AATE and baseline frameworks:

$$AE = \frac{\text{Revenue recovered}}{\text{Auditor hours or cost}} \qquad 4.10$$

Improvements in processing time, reduced manual workload, and optimized audit targeting contribute to measurable reductions in administrative costs. Scenario analysis allows evaluation under varying audit capacity constraints.

#### *4.4.3. Compliance Behavior and Elasticity*

AI-AATE may indirectly influence taxpayer compliance through perceived fairness and transparency. Compliance response is modeled using a simplified elasticity framework:

$$\Delta Compliance_i = \varepsilon_i . \Delta P_i \qquad \text{........} \qquad 4.11$$

where $\varepsilon_i$ is the compliance elasticity of taxpayer *i*, and $\Delta P_i$ is the change in perceived audit probability due to AI-driven selection. This accounts for behavioral adjustments without requiring real-world deployment, providing plausible estimates of system-wide effects.

#### *4.4.4. Cost–Benefit Synthesis*

Integrating revenue gains, administrative cost savings, and compliance elasticity, a net benefit metric is defined:

NetBenefit = \Delta R +\text{Cost Savings}- \text{Implementation & Maintenance Costs} 4.12

This framework allows sensitivity analysis, exploring conservative, moderate, and aggressive AI-AATE adoption scenarios. The model ensures that economic claims are quantifiable, evidence-based, and replicable.

### 4.5. Governance, Fairness, and Trust Metrics

To assess AI-AATE's institutional credibility, a set of governance and trust KPIs is integrated into the evaluation framework. These metrics quantify whether the system's design mechanisms (Ch.3) effectively produce transparent, accountable, and equitable audit outcomes.

Explainability Rate ( $EX$ ): Proportion of automated audit decisions accompanied by interpretable rationales accessible to auditors and oversight bodies.

$$EX = \frac{\text{Audits with full explanation}}{\text{Total audits}} \qquad 4.13$$

- Human Override Frequency ( $HO$ ): Fraction of AI recommendations modified or rejected by human auditors, indicating model conservatism and procedural checks.
- Equity Index ( $EI$ ): Measures disparities in audit selection across demographic, economic, or sectoral categories, flagging potential bias.
- Audit Traceability ( $AT$ ): Fraction of audit decisions with complete logged workflow and feedback integration, supporting retrospective review and regulatory compliance.

These governance metrics are simulated in parallel with operational and economic KPIs, enabling multi-dimensional assessment. They allow scenario-based testing of thresholds, fairness constraints, and feedback mechanisms, ensuring that AI-AATE achieves both efficiency and institutional legitimacy.

## 5. Conclusion and Implications

### 5.1. Summary of Contributions

This study presents the design and evaluation of AI-AATE, a secure, adaptive audit transparency engine for tax administration. Key contributions include:

- A modular AI architecture combining supervised and unsupervised learning with explainability and human-in-the-loop oversight.
- An evaluation framework linking operational KPIs, economic impact modeling, and governance metrics.
- Evidence from simulation and counterfactual modeling showing improvements in audit efficiency, revenue recovery, and fairness relative to traditional audit methods.

### 5.2. Implications and Recommendations

AI-AATE demonstrates that advanced, explainable, and auditable AI systems can transform tax administration by improving compliance while maintaining transparency and institutional trust. Practically, the framework provides policymakers and tax authorities with a blueprint for adaptive audit systems, highlighting the importance of integrating technical performance with governance and accountability mechanisms.

For future work, real-world deployment and empirical validation are recommended to confirm economic benefits, assess behavioral responses, and refine governance and transparency mechanisms. Further research could also explore applications beyond taxation, including customs, social benefits, and other areas of public finance.

In conclusion, AI-AATE exemplifies the potential of AI-driven, adaptive, and transparent audit systems to enhance revenue mobilization, reduce administrative inefficiencies, and strengthen overall economic prosperity, while ensuring fairness, accountability, and public trust. By combining technical rigor with institutional considerations, this framework represents a scalable model for the next generation of intelligent tax administration systems.

## Compliance with ethical standards

*Disclosure of conflict of interest*

No conflict of interest to be disclosed.

## References

[1] Alles, M., Brennan, G., Kogan, A., & Vasarhelyi, M. (2022). Continuous auditing and continuous monitoring: The state of the art and future directions. Accounting Horizons, 36(3), 1–20. https://doi.org/10.2308/acch-2022-003

[2] Azmi, N., Rahman, M., & Hassan, F. (2023). Transparency and trust in AI-driven decision-making: Challenges for public administration. Government Information Quarterly, 40(1), 101842. https://doi.org/10.1016/j.giq.2022.101842

[3] Azmi, R., Hussin, H., & Razak, A. A. (2023). Artificial intelligence adoption in public sector auditing: Risks, governance, and ethical challenges. Government Information Quarterly, 40(3), 101815. https://doi.org/10.1016/j.giq.2023.101815

[4] Bird, R. M., & Zolt, E. M. (2022). Tax compliance and fairness: Lessons from AI and digital auditing. National Tax Journal, 75(3), 523–550. https://doi.org/10.17310/ntj.2022.3.04

[5] Cobham, A., & Janský, P. (2023). Global tax losses and the corporate tax gap. World Development, 164, 106139. https://doi.org/10.1016/j.worlddev.2022.106139

[6] Edupuganti, P. (2024). Adaptive AI in auditing: Enhancing compliance and reducing operational risk. Journal of Accounting and Data Analytics, 8(1), 45–62. https://doi.org/10.1016/j.jada.2024.01.003

[7] Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design science in information systems research. MIS Quarterly, 28(1), 75–105. https://doi.org/10.2307/25148625

[8] IMF. (2023). Revenue mobilization and administrative efficiency: AI applications in taxation. International Monetary Fund. https://www.imf.org/en/Publications

[9] Iskandarova, N., KPMG, & Smith, L. (2022). Machine learning in auditing: Trends and applications. Accounting Horizons, 36(4), 67–85. https://doi.org/10.2308/acch-2022-003

[10] Kassa, T., & Taibi, D. (2023). Explainable artificial intelligence in public-sector decision-making: Opportunities and risks. AI & Society, 38(4), 1549–1565. https://doi.org/10.1007/s00146-022-01464-8

[11] OECD. (2021). Tax administration 2021: Digitalization, risk management, and audit strategies. Organisation for Economic Co-operation and Development. https://doi.org/10.1787/taxadmin-2021-en

[12] OECD. (2023). Tax administration 2023: Comparative information on OECD and other advanced and emerging economies. OECD Publishing.

[13] OECD. (2025). Predictive analytics in tax compliance: Case studies and best practices. Organisation for Economic Co-operation and Development. https://doi.org/10.1787/tax-analytics-2025-en

[14] Refining Public Policies with Machine Learning. (2024). AI applications for policy evaluation. PolicyTech Press.

[15] Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1135–1144. https://doi.org/10.1145/2939672.2939778

[16] Salmanov, R. (2024). AI and machine learning in auditing: Enhancing accuracy and efficiency. International Journal of Accounting and Information Systems, 29, 101–119. https://doi.org/10.1016/j.accinf.2024.101119

[17] Shehu, A., & Olukeye, O. (2024). AI, RPA, and NLP in tax reporting: Efficiency and compliance gains. Journal of Public Sector Innovation, 15(2), 35–52. https://doi.org/10.2139/jpsi.2024.15.2.35

[18] Transforming Auditing in the AI Era. (2025). Emerging technologies and audit transparency. Global Accounting Research Press.

[19] Zhang, Y., Xiong, F., Wang, Y., & Chen, Y. (2022). Machine learning–based tax risk management: A systematic review. Expert Systems with Applications, 198, 116789.